5-Hydroxymethylation marks a class of neuronal gene regulated by intragenic methylcytosine levels.

Genomics 104 (2014) 383–392

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

5-Hydroxymethylation marks a class of neuronal gene regulated by intragenic methylcytosine levels

Rachelle E. Irwin, Avinash Thakur, Karla M. O' Neill, Colum P. Walsh *

Centre for Molecular Biosciences, School of Biomedical Sciences, University of Ulster, Coleraine BT52 1SA, UK

ARTICLE INFO

Article history:
Received 4 July 2014
Accepted 18 August 2014
Available online 29 August 2014

Keywords:
Epigenetics
5′-Methylcytosine
5′-Hydroxymethylation
Neuronal function
Gametes
Development

ABSTRACT

We recently identified a class of neuronal gene inheriting high levels of intragenic methylation from the mother and maintaining this through later development. We show here that these genes are implicated in basic neuronal functions such as post-synaptic signalling, rather than neuronal development and inherit high levels of 5mC, but not 5hmC, from the mother. 5mC is distributed across the gene body and appears to facilitate transcription, as transcription is reduced in DNA methyltransferase I (*Dnmt1*) knockout embryonic stem cells as well as in fibroblasts treated with a methyltransferase inhibitor. However in adult brain, transcription is more closely associated with a gain in 5hmC, which occurs without a measurable loss of 5mC. These findings add to growing evidence that there may be a role for 5mC in promoting transcription as well as its classical role in gene silencing.

1. Introduction

DNA methylation in mammals usually refers to the modification of cytosine by the addition of a methyl group to the 5′ position (5mC), usually at the CpG dinucleotide. However, study of the Ten-Eleven Translocation (TET) proteins and advances in sequencing approaches have revealed that TETs can convert 5mC to 5-hydroxymethylcytosine (5hmC), which is widely found in brain and other CNS components (see review articles by the Tang and Pfeifer labs, this issue). TETs can catalyse the formation of the further derivatives 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), which appear to be turned over more rapidly, at which point the modified base possibly is removed from the DNA by the base excision repair enzyme Thymidine DNA Glycosylase (TDG) [1–3].

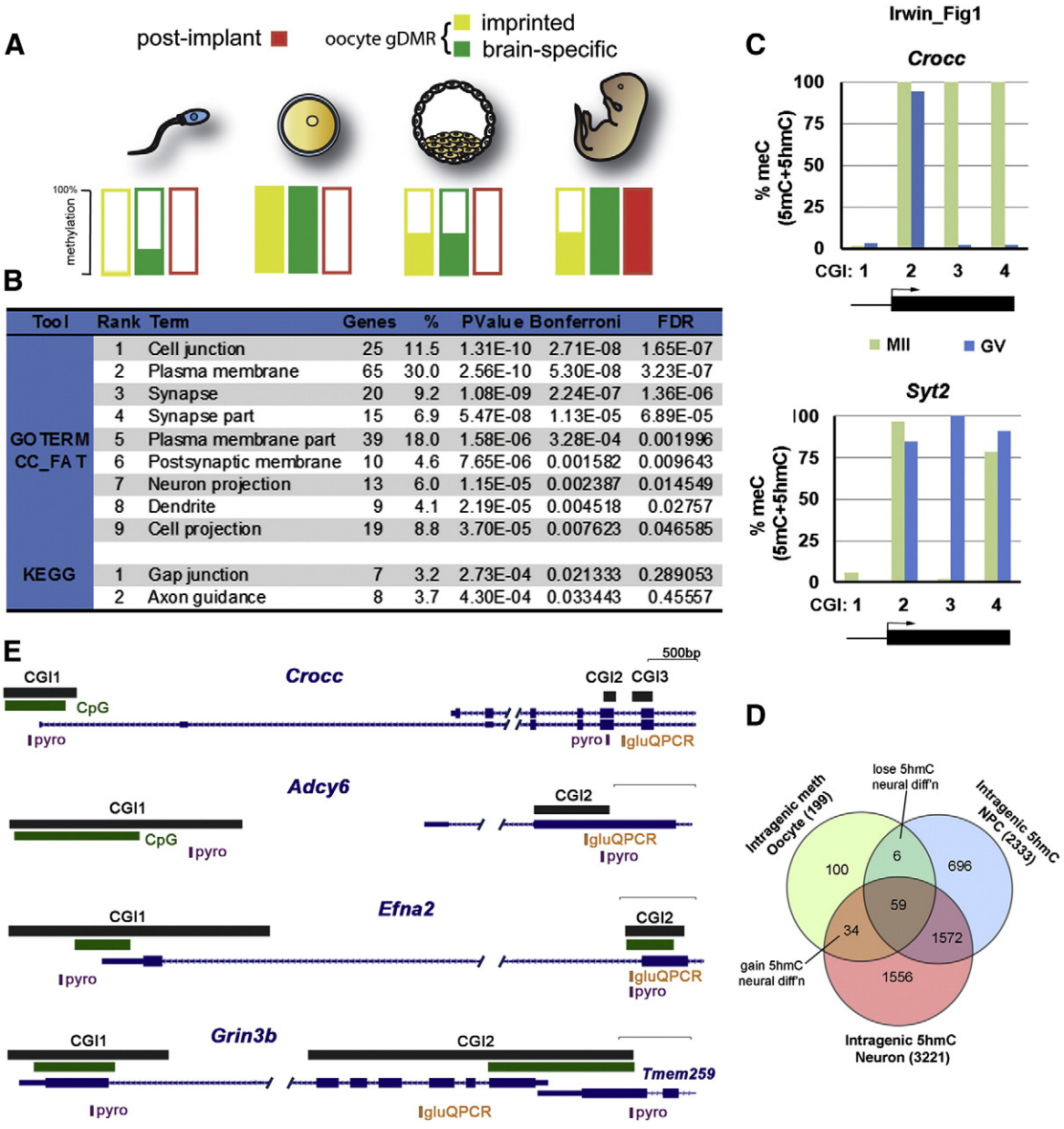
5hmC and further products can only be formed from 5mC, as in the absence of this base none of the former bases are found [4], suggesting that TET-mediated oxidation is the main pathway for active demethylation in mammals. Consistent with this, loss of TET3 prevents the demethylation of DNA in the paternal pronucleus post-fertilisation [5]. In contrast, the relatively steady levels of 5hmC in neural tissues, which have very low turnover rates, suggest that it may be playing a different role there. A number of studies have found 5hmC to be enriched in gene bodies in neural precursor cells (NPC), and to be positively correlated with transcription [6–8]. Loss of TET1 compromised the self-renewal capacity of NPC and led to decreased transcription of a number of genes involved in progenitor proliferation [6]. Likewise, it has recently been shown that loss of intragenic 5mC led to a decrease, rather than an increase, in transcription levels in embryonic stem cells (ESC) lacking DNMT3L [9], suggesting that this type of methylation may be needed to achieve high levels of transcription. Another study has also shown that transcription of a class of brain-specific genes most highly expressed in NPC was diminished in the absence of the de novo enzyme DNMT3A [10].

Roles for DNA methylation in promoting transcription are in sharp contrast to its “classical” role in repressing transcription when found at promoter CpG islands (CGI). CGI have been identified by bioinformatic means but more recently, biochemical assays based on binding of the CXXC domain protein CFP1 (CXXC1) have been used to identify CGI conserved between mouse and human, almost half of which were not detected bioinformatically [11,12]. Many CFP1-defined CGI are located away from annotated promoters in intra- or intergenic locations, so-called “orphan” CGI [13].

We recently identified a class of neuronal gene based on their methylation ontogeny (16). These genes appeared to become methylated at specific intragenic CGI in the oocyte, then retain this methylation through later development (Fig. 1A). Methylation was significantly lower or absent in sperm, then the remaining methylation was essentially lost by the blastocyst stage: following implantation, the paternal allele became de novo methylated leading to almost complete methylation of both alleles in adult tissues. These genes thus resembled imprinted gene ICR in their early methylation ontogeny, but differed from them in their inability to retain a methylation-free paternal allele post-implantation (Fig. 1A).

* Corresponding author.
E-mail address: cp.walsh@ulster.ac.uk (C.P. Walsh).

http://dx.doi.org/10.1016/j.ygeno.2014.08.013
0888-7543/



Tool	Rank	Term	Genes	%	PValue	Bonferroni	FDR
GO TERM CC_FAT	1	Cell junction	25	11.5	1.31E-10	2.71E-08	1.65E-07
	2	Plasma membrane	65	30.0	2.56E-10	5.30E-08	3.23E-07
	3	Synapse	20	9.2	1.08E-09	2.24E-07	1.36E-06
	4	Synapse part	15	6.9	5.47E-08	1.13E-05	6.89E-05
	5	Plasma membrane part	39	18.0	1.58E-06	3.28E-04	0.001996
	6	Postsynaptic membrane	10	4.6	7.65E-06	0.001582	0.009643
	7	Neuron projection	13	6.0	1.15E-05	0.002387	0.014549
	8	Dendrite	9	4.1	2.19E-05	0.004518	0.02757
	9	Cell projection	19	8.8	3.70E-05	0.007623	0.046585
KEGG	1	Gap junction	7	3.2	2.73E-04	0.021333	0.289053
	2	Axon guidance	8	3.7	4.30E-04	0.033443	0.45557

Fig. 1. Intragenic methylation is established in the oocyte for a novel class of brain-specific gene. (A) Many genes only become methylated in the post-implantation wave of de novo methylation (red). Some brain-specific genes have gametic differentially methylated regions (gDMR), inheriting high levels of methylation from oocytes, and are in this way comparable to imprinted genes. Unlike imprinted genes, brain genes are not protected from methylation post-implantation. % meC (5mC + 5hmC) represents total % methylated cytosine, as a sum of 5mC and 5hmC detected by pyrosequencing. (B) Analysis of the brain gDMR genes by DAVID indicates enrichment for general neuronal functions rather than neurodevelopmental terms. *P* value, Bonferroni-corrected *P* value and false discovery rate (FDR) are indicated. (C) Structure and methylation at two representative brain genes with multiple CGI, numbered from 1 (promoter) to 4 (3′). Combining data from reduced representation bisulfite sequencing (RRBS) in meiosis II (MII) oocytes [17] with whole bisulfitome amplified DNA sequencing on germinal vesicle (GV) oocytes [18] indicates methylation is not targeted to any one internal CGI. Promoter CGI are generally unmethylated (CGI1). (D) Many of the brain genes carrying gDMR from oocyte were independently identified as having intragenic 5-hydroxymethylation (5hmC) in either neurons or neuronal precursor cells (NPC) or both [7]. (E) Structure of the four brain loci studied in more detail, indicating the positions of the CFP1-identified CGI (black, top line), classical CGI (green, second line) and probes used for combined 5mC/5hmC analysis by bisulfite pyrosequencing (pyro) and for 5hmC-specific assays (gluQPCR).

A number of questions remain unresolved with respect to this gene class and their methylation. In particular, we wished to 1) determine if there were any common features with respect to the function of these genes, 2) determine whether methylation was confined solely to one intragenic CGI, 3) determine what fraction, if any, of the cytosines appearing in bisulfite-converted samples are in fact 5hmC rather than 5mC, 4) determine in more detail when and where methylation of the paternal allele occurs post-implantation and 5) determine if there was any functional requirement for the methylation by removing it and examining transcriptional response.

We show here that this neuronal gene class showing high levels of intragenic methylation inherited from the mother are implicated in basic neuronal functions such as post-synaptic signalling, rather than developmental processes. These genes inherit high levels of 5mC, but not 5hmC, largely from the mother, which may facilitate a poised transcriptional state in ESC and cells of the blastocyst inner cell mass (ICM). De novo methylation during differentiation further primes them for high-level transcription, but this is only realised in neuronal tissues, where appropriate transcription factors can be found, and is associated with high levels of 5hmC. In contrast, methylation in non-neuronal tissues consists almost exclusively of 5mC. Using pharmacological as well as genetic approaches we demonstrate that while 5mC alone is insufficient to allow high-level transcription, its loss causes reproducible decreases in transcription levels in all tissues examined.

2. Results

2.1. Bioinformatic analysis of methylation state of a novel class of neuronal genes

To determine what common features the novel neuronal gene class might have, we looked for common ontologies among the genes using the Functional Annotation and KEGG tools in DAVID [14,15]. The former identified significant enrichment of terms related to synapse ($P = 1.08 \times 10^{-9}$), post-synaptic membrane ($P = 7.65 \times 10^{-6}$) and dendrite ($P = 2.19 \times 10^{-5}$) among others (Fig. 1B), while the latter highlighted, at lower significance (both $P > 1 \times 10^{-4}$) gap junction and axon guidance terms as over-represented in the group.

Our previous analyses had indicated that CGI being flagged as methylated by genome-wide analyses in oocyte were all intragenic, often at a specific CGI, while the promoter CGI for these genes, when present, were unmethylated in this cell type [16]. Combining data from two extant studies [17,18] and examining genes with multiple intragenic CGI confirms that promoter CGI appear to be unmethylated, but suggests that all intragenic CGI, and most likely the entire body of the gene, are methylated: two examples *Crocc* and *Syt2* are shown (Fig. 1C). However significant differences between the studies in terms of results for individual CGI underline the need for gene-specific assays to confirm which pattern is correct at loci showing discrepancies. Notably, both of the studies employed bisulfite-treatment followed by sequencing, which cannot distinguish between 5mC and 5hmC. This is significant since half of the genes in our group are flagged as having intragenic, but not promoter, 5hmC enrichment in brain in a recent study [7] (99/199 in areas of overlap—Fig. 1D). Notably a large fraction of these are already enriched in 5hmC in neural precursor cells (NPC) (65/99—Fig. 1D), suggesting that hydroxylation may be occurring earlier in development, or potentially even be inherited from the gametes. To further explore these questions we designed a set of assays to investigate the levels of 5mC and 5hmC at representative genes from our neuronal gene set (Fig. 1E).

2.2. High levels of 5mC but not 5hmC in the neuronal gene set in gametes

We first confirmed that methylation in the oocyte was confined to the intragenic CGI: examination of the total methylation levels (5mC + 5hmC) by bisulfite treatment followed by pyrosequencing indicated that methylation was below 25% at promoter CGI (Fig. 2A), as indicated in the genome-wide assays. We could also confirm high (>50%) methylation at the intragenic CGI examined (Fig. 2A), although there was more variability here. On the whole pyrosequencing results correlated well with genome-wide assays, although the average methylation level of the 8 CpG in CGI2 of *Crocc* was substantially lower using the former technique. While intragenic CGI show significantly higher methylation levels generally in oocyte than sperm as we previously reported, the downstream CGI are also more methylated than the promoter in sperm (Fig. 2B).

To determine how much, if any, of the methylation detected using bisulfite-dependent techniques is actually 5hmC and not 5mC, we used Glu-qPCR, which takes advantage of the fact that the former base alone can be glucosylated [19]. Using this we could detect relatively low levels of 5hmC in oocytes, ranging from 1.3% (*Crocc*) to ~3.6% (*Grin3b*) (Fig. 2C). In contrast, levels of 5hmC at *Grin3b* in brain using this assay were ~30% (Fig. 2C—positive control). Hydroxymethylation levels at other gene classes (imprinted, testis-specific) which show oocyte-specific methylation were slightly higher (Fig. 2C), but all were <10% and differences were not statistically significant. Likewise, 5hmC levels in sperm were generally very low, with *Grin3b* being the highest at 8.5% (Fig. 2D), but all other genes being uniformly lower than 2%.

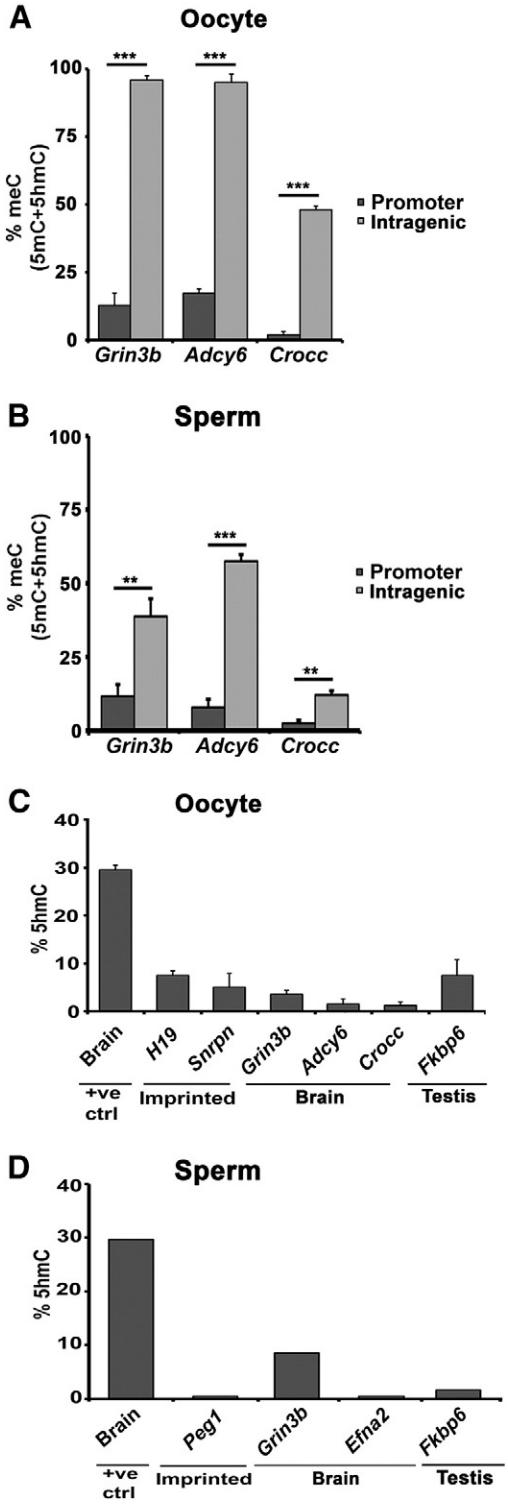


Fig. 2. Intragenic gametic methylation exists in the form of 5mC, not 5hmC, for the identified brain genes. (A) Methylation is enriched at intragenic CGI in MII oocytes, with the promoter CGI remaining unmethylated. (B) Increased methylation at intragenic CGI in comparison with promoter CGI in sperm. (C) The level of 5hmC in oocytes is ubiquitously low for the gene classes analysed, in comparison with the brain (+ve ctrl: positive control). (D) Sperm is also largely devoid of 5hmC in the gene classes examined.

2.3. High but variable methylation levels in adult tissues

Comparison of methylation levels by pyrosequencing in oocyte, ESC (which model the ICM of blastocyst) and soma confirmed our previous

observations that the brain group become fully methylated during post-implantation development, presumably through de novo methylation of the paternal allele (Fig. 3A). In contrast, the ICR at *Peg1 (Mest)* remains unmethylated (Fig. 3A), as expected of an imprinted gene. Previous analysis of a small number of adult mouse tissues [16] had suggested that methylation was uniformly high for the brain genes. However, extension of this analysis to additional tissues (Fig. 3B) showed considerable variability existed, with methylation levels as low as 25–35% in intestine (*Adcy6*) and ovary (*Crocc*).

We considered that the low levels of methylation as observed in some adult tissues may reflect loss of methylation in expressing tissues and that the brain gene group might have members that are more widely transcribed than just in neural tissues. To examine this we carried out RT-qPCR in different adult tissues: our results (Fig. 3C) indicated that the majority of the brain genes examined were highly specific to brain tissue, with the exception of *Adcy6* which shows transcription levels in non-neural tissues of almost 50% the level seen in brain (Fig. 3C). In adult tissue there was no clear correlation between levels of total DNA

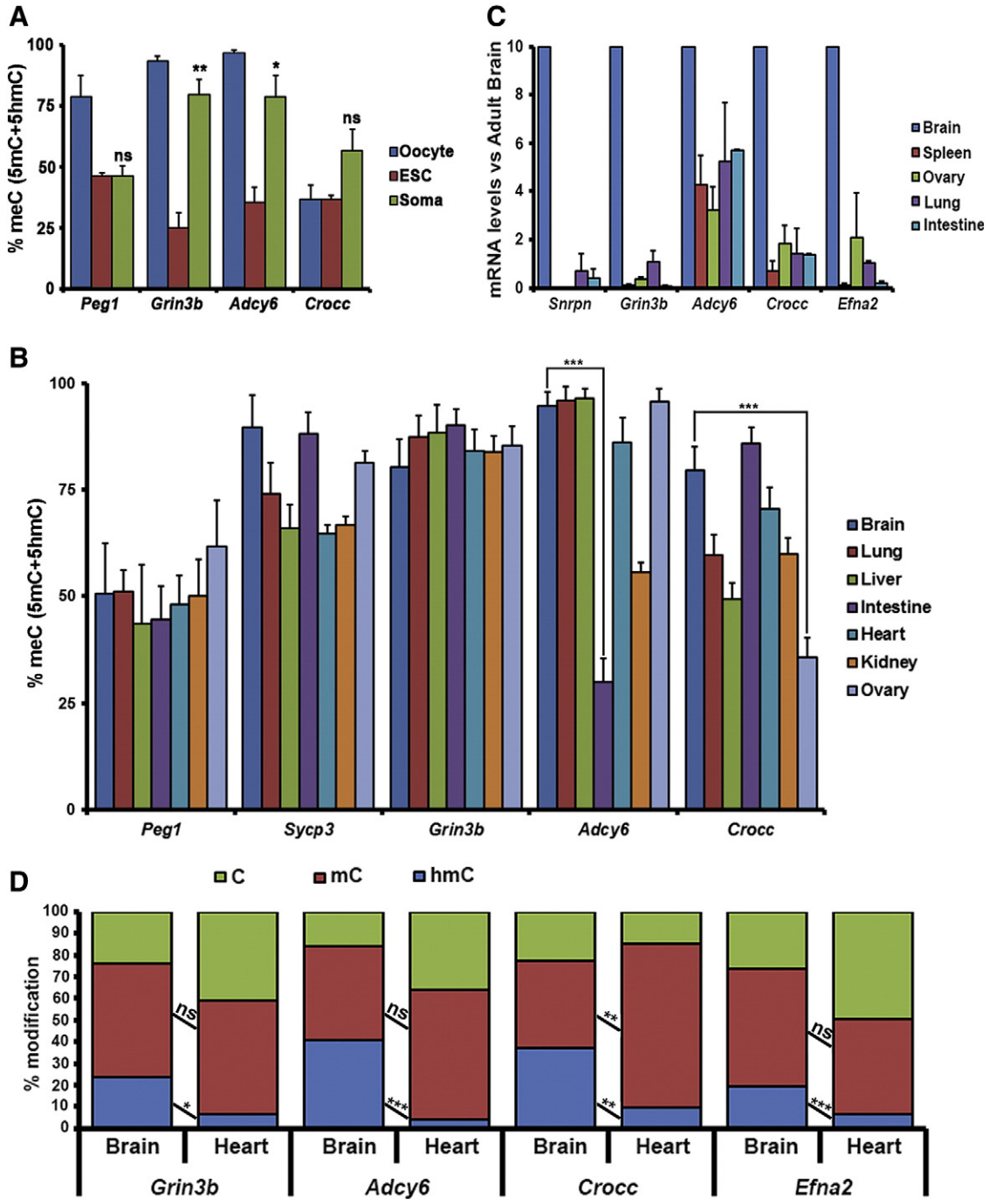


Fig. 3. Variability in total methylation levels between post-implantation tissues does not correlate with expression for brain genes examined. (A) Methylation profile during development from MII oocyte, J1 (WT) ESC and soma (average of all adult tissues) indicates maternal inheritance of methylation at the brain genes examined, which is retained in the ESC and then lost post-implantation via de novo methylation of the paternal allele. Only *Peg1* retains this parent-of-origin mark, and thus is protected from methylation in the soma. (B) The methylation levels in individual adult tissues show variability. (C) Fold expression levels, relative to adult brain, highlight predominant transcription in the adult brain. Levels in brain were normalised to 10 (D) Absolute cytosine modification levels (C, mC, hmC) at individual CpG sites measured via Glu-qPCR in brain and heart tissues. Although similar total methylation levels exist, the brain is enriched for 5hmC in comparison with heart (4/4 genes), with no significant difference in 5mC (3/4).

methylation (5mC + 5hmC, as measured by bisulfite pyrosequencing) and transcription levels, as a number of non-expressing tissues had little 5mC + 5hmC (e.g. *Crocc* in ovary, Fig. 3C) while other tissues with high mRNA levels had high methylation (e.g. *Grin3b* in brain).

Variation in the proportions of 5hmC to 5mC is one possible explanation for the lack of correlation between transcription and total methylation levels when using bisulfite-based techniques. In order to distinguish 5hmC and 5mC but measure them in the same assay, we used a second type of Glu-qPCR which can quantitate the relative levels of unmodified cytosine as well as 5mC and 5hmC. Examination of brain, where the target genes are transcribed at high levels (Fig. 3D), showed that as much as 40% of the methylation at individual CpG sites assayed was 5hmC rather than 5mC (Fig. 3D). In contrast, other tissues such as heart showed levels of 5hmC which were significantly, or very significantly, lower than in brain (Fig. 3D). On the other hand, levels of 5mC showed no significant difference between brain and heart at 3/4 loci, indicating that adult brain is distinguished not by a decrease in 5mC, but by an increase in 5hmC (Fig. 3D).

2.4. De novo methylation occurs early during differentiation and correlates with increased transcription of the brain gene group

To determine if de novo methylation is occurring during the early stages of differentiation or whether it may be occurring later in a more tissue-specific fashion, we took advantage of a previously-described in vitro differentiation system [20]. In this system ESC were differentiated initially into primary embryoid bodies, which resemble the blastocyst, then into secondary embryoid bodies which contain populations of neural precursor cells (NPC) [21] before being specialised into haematopoietic stem cells (Fig. 4A). As controls, we confirmed that in this system the imprinted gene *Peg1* remains stable in the 40–60% range seen in normal tissues [22], as expected of an imprinted gene (Fig. 4B), while the germline-specific gene *Sycp3* undergoes de novo methylation from 25% to >80%, as seen in vivo [16]. The brain gene group were also found to undergo de novo methylation as seen in vivo (Figs. 3A,B) leading to >75% methylation at *Grin3b*, *Adcy6* and *Crocc* and a gain in methylation of >30% at all four loci examined. These results were confirmed using combined bisulfite modification and restriction analysis (COBRA) (data not shown).

As expected from in vivo studies and our previous characterisation of this system, transcription of the germline gene *Sycp3* decreased significantly as methylation at the promoter CGI increased during differentiation (Fig. 4C). In contrast methylation at the intragenic CGI of the brain genes correlated instead with increasing transcription levels during the initial stages (Fig. 4C), before transcription drops in the more specialised HSC (most likely due to the lack of required transcription factors). This trend was especially clear for *Adcy6* (Fig. 4D). Transcription of all the brain genes was highest in secondary embryoid body, which contains the most NPC. Likewise, there was a trend to higher transcription levels in foetal brain over adult (Fig. 4E), although this was not statistically significant except for *Adcy6*. Examination of 5hmC and 5mC levels (Fig. 4F) showed that 1) combined methylation (5hmC + 5mC) levels increased from foetal to adult brain; 2) 5mC levels were higher in foetal brain and 3) in adult a greater fraction of the total methylation consisted of 5hmC.

2.5. Depleting 5mC results in decreased transcription of the brain gene group

The correlation between 5mC levels and transcription suggested that this modification may somehow facilitate transcription across the brain genes. To test this, we first examined ESC deficient in the maintenance DNA methyltransferase enzyme DNMT1. This confirmed that when methylation was significantly reduced (Fig. 5A) mRNA levels showed reproducible decreases for all four brain genes examined (Fig. 5B). This is in contrast to a control imprinted gene which is unaffected (Fig. 5B) and to testis genes with methylated promoter CGI, which we have previously shown to be upregulated [16]. A similar effect was seen in *Dnmt3a* $^{-/-}$*Dnmt3b* $^{-/-}$ double mutant ESC, though to a lesser extent as expected from their higher residual methylation level (data not shown). To confirm and extend this, we examined the effects of the pan-methyltransferase inhibitor 5′-aza-2′deoxyctidine (Aza) in adult fibroblasts. While the drug was less effective at reducing methylation, levels did decrease significantly at many loci (Fig. 5C). Treatment resulted in small but significant and reproducible decreases in transcription for 3/4 brain genes (Fig. 5D). In contrast, Aza treatment resulted in derepression of *Sycp3*, normally repressed by methylation of its promoter CGI post-implantation (Fig. 5D).

2.6. Conservation of methylation type and location at brain homologs in humans

An important consideration in terms of determining functionality is the degree of conservation between different species. Many of the neuronal genes described above have human homologs: examination of these loci showed that intragenic CGI were also often conserved. Data from the UCSF brain DNA methylation study [12] typically shows conservation of 5mC distribution as well, with an unmethylated promoter CGI but methylated intragenic CGI and gene body, as typified by *ADCY6* (Fig. 6A). We extended these observations to primary fibroblasts by generating genome-wide methylation data using the 450k beadarray system: this data confirms that upstream promoter CGI for the brain genes tend to be unmethylated (PCS—Fig. 6A, bar graph), while downstream CGI in the gene body tend to be heavily methylated. Recent work by Pfeifer et al. compared 5mC and 5hmC frequency genome-wide using DNA immunoprecipitation-based techniques (DIP) [23]. Of 53 human homologs of our mouse genes that appeared in that dataset, 18/53 showed high levels of both 5mC and 5hmC intragenically from their data: this fraction rose to 26/53 (49%) when combined with data on 5mC distribution from the UCSF project and our own 450K dataset. Since the DIP datasets have limited coverage compared to resequencing-based approaches, this suggests that, conservatively, at least half of all brain genes showing high levels of intragenic 5mC are also enriched for intragenic 5hmC, as in mouse. In contrast, testis genes showed low levels of intragenic 5hmC and 5mC, were transcriptionally silent in brain and had high 5mC but little 5hmC at their promoters [16,23] (also UCSF and 450K, data not shown).

To examine whether intragenic 5mC may play a similar role in human at these loci, we used the hTERT-immortalised normal human fibroblast cell line hTERT-1604, which we have previously found to have a similar methylation profile to primary cells using locus-specific approaches [24]. Initial observations from 450 K analyses of independent samples suggest that these cells do indeed have a similar profile to primary fibroblasts at the majority of CpGs examined (see for example data for *ADCY6* in Fig. 6A bar graph). We confirmed intragenic methylation in these cells by pyrosequencing at a number of loci (Fig. 6B, cells—Aza). As for most other cultured adult human cell lines [25], levels of 5hmC were very low at all sites assayed (data not shown). Treatment of these cells with Aza gave a small decrease in transcription at 2/3 brain genes examined, while the imprinted gene SNRPN was unaffected (Fig. 6C). The conservation of position, methylation status and effects on transcription of these CGI in human further supports their important role in regulation of these genes.

3. Discussion

The class of brain-specific genes described here were initially identified by us based on their higher levels of methylation in oocyte than sperm, their dependence on DNMT3L for methylation establishment and the transient nature of the DMR, which is erased by de novo methylation of the paternal allele post-implantation [16]. In that study, we

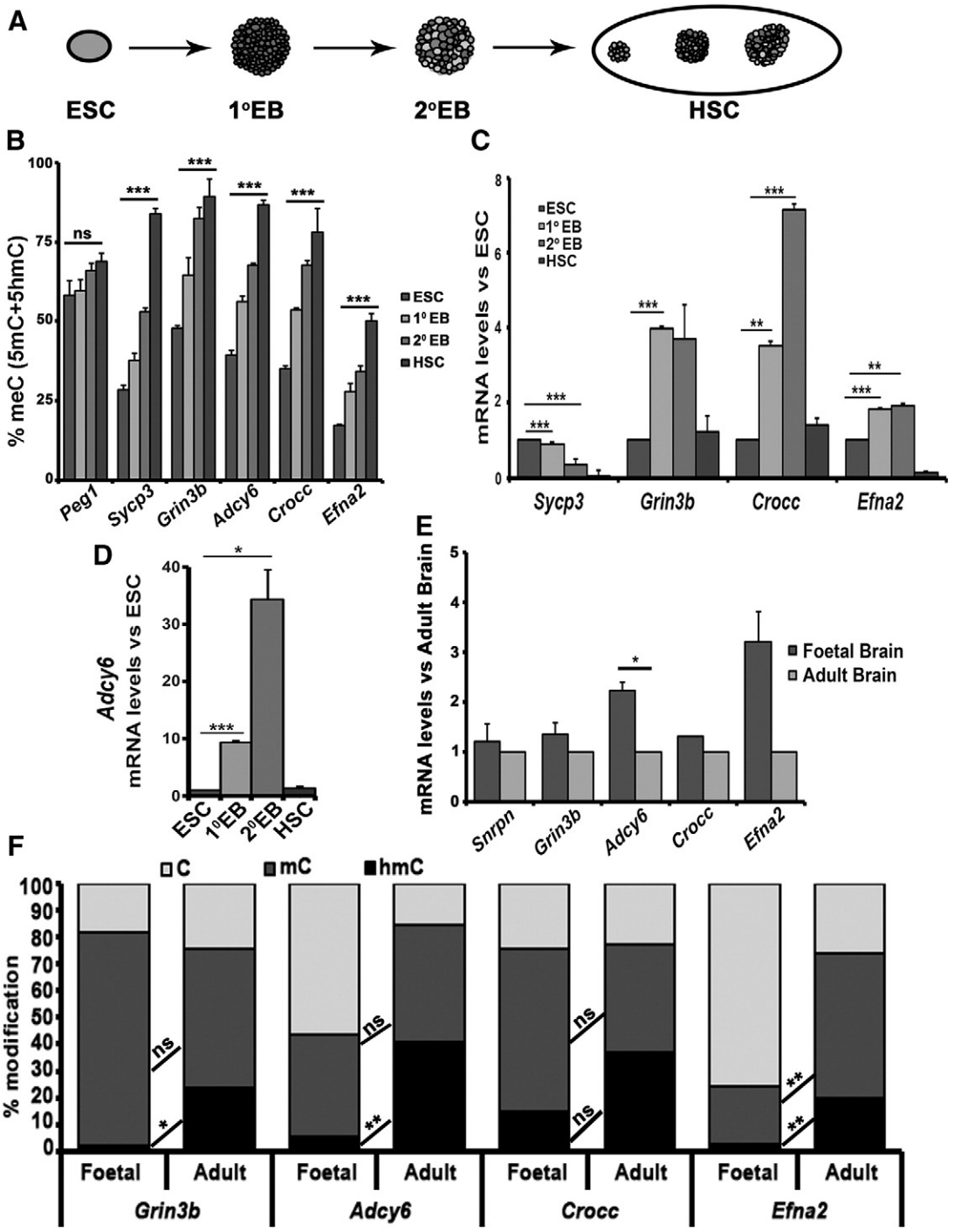


Fig. 4. De novo methylation occurs early during differentiation and correlates with increased transcription of the brain genes. (A) Schematic of in vitro ESC differentiation experiment, from undifferentiated ESC to primary embryoid body (1°EB), secondary embryoid body (2°EB) and haemopoietic stem cell (HSC). (B) Methylation of brain genes increases progressively during in vitro differentiation. (C) Expression levels increase with ESC differentiation, peaking in 2°EB which contains many NPC, before declining in HSC. *Sycp3* is a control and shows the opposite pattern, as previously reported (D) *Adcy6* is upregulated similarly to genes in (C) in 2°EB, relative to ESC. (E) *Adcy6* and *Efna2* are upregulated in foetal (~e15.5) brain relative to adult brain. (F) Absolute cytosine modification levels (C, mC, hmC) at individual CpG sites determined via glu-qPCR indicate significantly higher hmC levels (3/4), but not mC levels (3/4), in adult brain.

identified a second class of genes sharing these characteristics. This second class differed in that 1) they were largely testis-specific in their transcription rather than brain-specific; 2) methylation was present at the promoter rather than intragenically and 3) methylation repressed the genes. Here we further characterise the brain genes and show that methylation for these genes encompasses the whole gene body, while the promoter is generally methylation-free, and we provide evidence that the methylation facilitates transcription of the gene rather than repression, since in both mouse and human, decreasing methylation reduces rather than increases levels of mRNA.

The recent discovery of 5hmC and the finding that bisulfite-dependent methods do not distinguish 5mC from 5hmC led us to analyse 5hmC levels at the brain genes. Our results indicate that 5hmC levels are highest for the novel genes in the adult brain, while levels in other adult tissues and in gametes are much lower, suggesting that 5hmC does not contribute a substantial proportion of the inherited

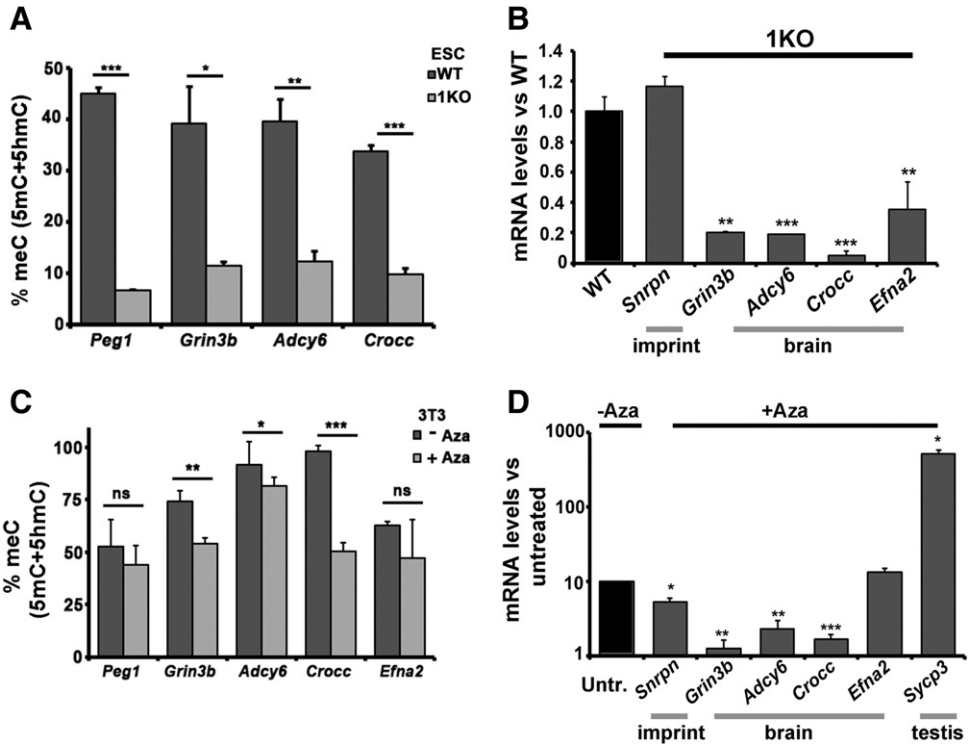


Fig. 5. Depletion of methylation correlates with reduced transcription of brain genes. (A) Intragenic methylation of brain genes decreases in *Dnmt1*-depleted ESCs (1KO). (B) Removal of methylation results in down-regulation of brain genes in ESCs. (C) Methylation levels in adult mouse fibroblasts (3T3) before (−) and after (+) Aza treatment. (D) Aza treatment causes transcriptional repression of brain genes assayed relative to untreated (Untr.) controls. *Sycp3* is a control and shows the opposite pattern, as expected.

methylation at these loci but instead is generated during later neural development. How this fits with 5hmC being a transient product in the active demethylation of 5mC is not yet clear, but our findings are consistent with work by Pfeifer, Ecker et al. [8,23], who describe high levels of 5hmC co-localised with high levels of 5mC in adult brain, suggesting a relatively stable mark (see also review by Pfeifer et al. in this issue).

The reason why genes involved in steady-state neuronal function should have particularly high levels of intragenic 5mC in oocyte and early embryo is not currently clear. Consistent with our findings described here, recent papers by others [9,10] have also uncovered a previously unappreciated role for intragenic 5mC in promoting gene transcription, since loss of DNMT3L in ESC, or DNMT3A in brain, led to decreased mRNA levels for many genes normally transcribed in brain. The genes identified in those studies do not overlap significantly with those described here. Consistent with this, neurodevelopmentally important genes were largely absent from our group, while featuring strongly elsewhere. It is possible that a need for high-level transcription of the genes we have identified in early neurulating embryos may explain the high methylation levels of these genes in oocyte and early embryo. Recently, evidence has been emerging that the presence of read-through RNA, enabled by a more open chromatin structure, may lead to transcriptionally-coupled DNA methylation (reviewed in [26]). In support of this, the brain genes show leaky expression in both mouse and human fibroblasts, which facilitated measurement of down-regulation. Intragenic 5mC and read-through transcription are not a default state, as many other genes such as the testis genes show low intragenic methylation and have undetectable transcript levels in adult tissues.

There may be a loss of 5hmC as well as 5mC in the *Dnmt1* KO mouse ESC, since the hydroxymethylated form appears to be derived solely from the methylated form [4]. However, with the exception of ESC, cultured cell lines contain little or no 5hmC [25], so the repression of transcription of our novel gene class which we see in mouse and human fibroblast lines must be due to loss of 5mC. Whether the intragenic 5hmC is absolutely required for the high-level transcription seen in adult brain for this gene class is currently unclear. Due to the anatomical complexity of the brain, increased 5hmC in adult brain overall may mask decreases in specific structures, or vice versa. Further analyses using tissue-specific knockouts of the TET enzymes in the future may help to clarify the role of 5hmC at these genes.

4. Materials and methods

4.1. Genome-wide methylation analysis and bioinformatics

For mouse, bioinformatic analysis was performed on published genome-wide methylation datasets [7,12,17,18,23] using publicly available tools. Gene ontology analysis, including statistical significance, was carried out using DAVID [14,15]. For human, DNA was isolated from the hTERT-immortilised lung fibroblast (hTERT-1604) cell line using the phenol:chloroform:isoamyl alcohol (25:24:1—pH 8, Sigma-Aldrich, Dorset, UK) extraction method (below). DNA from the foreskin fibroblast cell line (PCS) was a kind gift from Dr. Fulvio Gandolfi. Purity and integrity of the DNA preparations were confirmed by 1% agarose gel electrophoresis and absorbance measurements at 260/280 and 260/230, prior to quantification using the Picogreen fluorescent assay (Invitrogen, Paisley, UK) as per as per manufacturer's instructions. In total, 500 ng of high quality bisulfite converted (Zymo Research Corporation, California, USA) DNA was analysed on the Illumina 450 k beadarray which was then imaged using an Illumina iScan (Cambridge Genomic Services, Cambridge). Output data were filtered for the presence of SNPs and other QC factors and remaining CpG beta-values were converted to % methylation and plotted against genomic loci using GALAXY software [27].

4.2. Statistical analysis

Experiments were carried out in triplicate, excepting the Glu-qPCR for oocyte and sperm, which were carried out in duplicate.

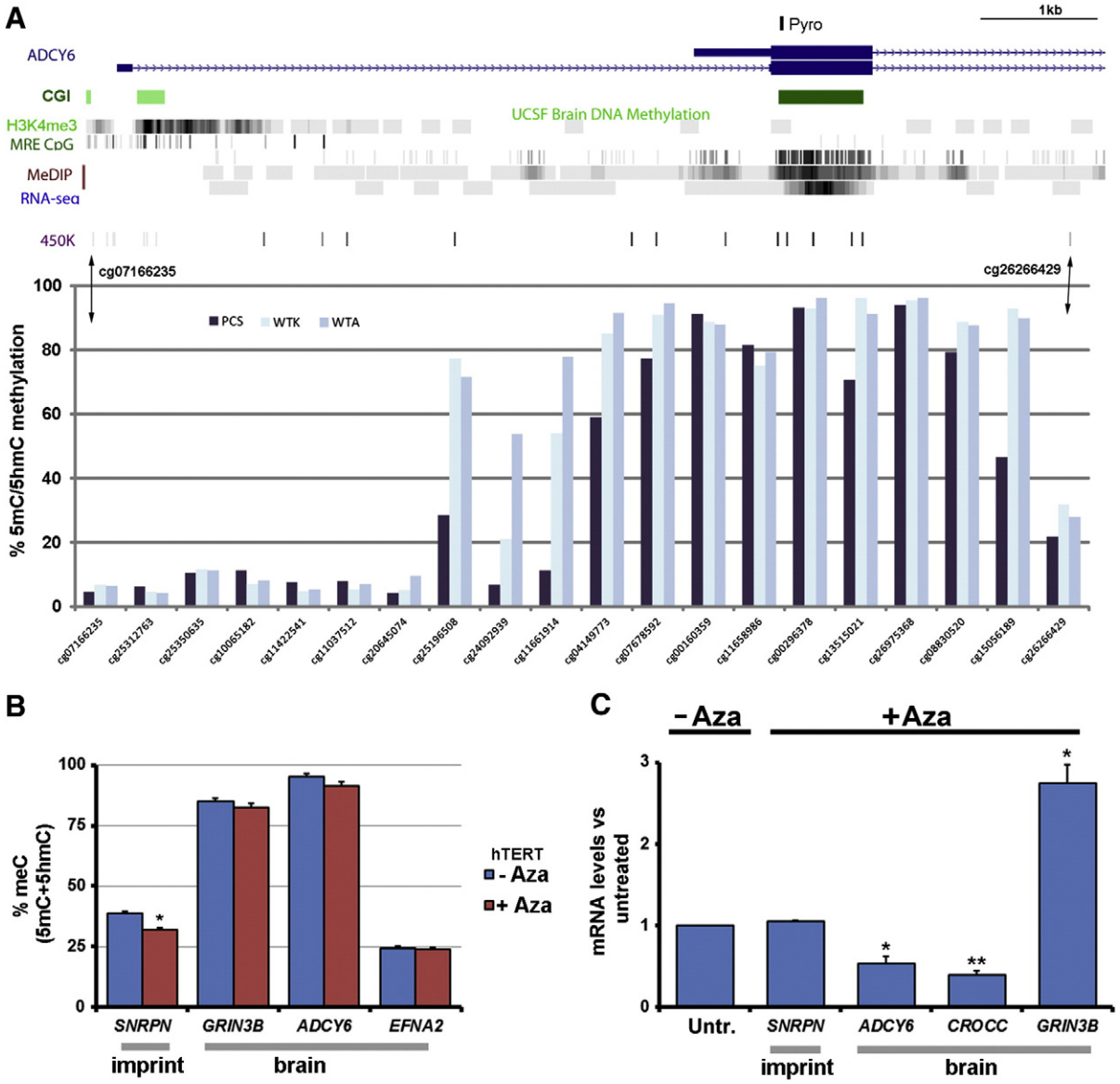


Fig. 6. Evidence for conservation of brain gene structure, methylation profile and response to demethylation in human. (A) Representative gene homolog in human showing conserved location of the promoter and intragenic CGI. Tracks from the UCSF brain methylation study [12] show evidence for an unmethylated promoter (MRE CpG signal) consistent with the presence of trimethylation of histone 3 lysine 4 (H3K4me3). In contrast, the intragenic CGI is methylated (MeDIP tracks) and the gene itself is actively transcribed (RNA-seq track). Analysis of genome-wide methylation in both primary (PCS) and two samples of hTERT1604 fibroblasts (Wild type A and K: WTA and WTK) confirmed low levels of methylation at the promoter CGI and high levels throughout the gene in both cell types (bar graph: location of indicative CpG assayed is shown). (B) Methylation levels in hTERT1604 cells before (−) and after (+) Aza treatment. (C) Aza treatment causes transcriptional repression of 2/3 brain homologs assayed relative to untreated (Untr.) controls. The imprinted *SNRPN* gene was unaffected.

Each set of experiments included at least one biological repeat. Statistical analysis of results was carried out in PRISM, EXCEL or GRAPHPAD; significance values throughout were based on Student's *t*-test, and error bars represent standard error of the mean (SEM) on all graphs. Significance values are in some cases abbreviated as $^{*}p < 0.05$; $^{**}p < 0.01$; and $^{***}p < 0.001$.

4.3. Cell culture and treatments

hTERT-1604 and NIH 3T3 cells were cultured as detailed in Rutledge and Thakur et al. (2014).The mESCs, comprising WT J1, *Dnmt1* KO and *Dnmt3a*/3b KO, were a kind gift from Dr. M. Okano (RIKEN Center for Developmental Biology, Kobe, Japan). ESCs were cultured as previously described [16]. Material from in vitro differentiated ESCs was a kind gift from Dr. Helen Wheadon (Paul O'Gorman Leukaemia Research Centre, Glasgow, UK). For pharmacological inhibition of DNMT1 during culture, 1×10^5 hTERT-1604 or NIH 3T3 cells were treated with 5-aza-2′-deoxycytidine (5′aza-dC; Sigma-Aldrich, Dorset, UK), at a final concentration of 1 μM, on 3 consecutive days.

4.4. Animal work

Tissues of interest derived from outbred TO mice (Harlan, Huntingdon, UK). Oocyte and sperm collection was carried out as previously described [28,29]. All animal work was carried out under licensing from the UK Home Office.

4.5. DNA extraction and bisulfite treatment

DNA extraction from oocytes and sperm was performed as previously described [28,29]. DNA was extracted from cell pellets and tissues using the phenol:chloroform:isoamyl alcohol (25:24:1—pH8,

Sigma-Aldrich) extraction method, involving overnight rotation–incubation at 55 °C in lysis buffer (50 mM Tris pH8, 0.1 M EDTA, 0.5% SDS, 0.2 mg/ml Proteinase K). Approximately 150–200 oocytes or 500 ng of DNA was bisulfite converted using the EpiTect Bisulfite Kit (Qiagen, Crawley, UK).

4.6. RNA extraction, cDNA synthesis and RT-qPCR

RNA was extracted using the RNeasy Kit (Qiagen) according to the manufacturer's instructions. cDNA was synthesised from 250 to 500 ng of RNA, combined in reaction with 0.5 μg random primers (Roche), 0.5 μM dNTPs, 40U RNaseOUT (Invitrogen), 1x Buffer RT and 200 U RevertAid reverse transcriptase (both Fermentas), and made up to a final volume of 20 μl using RNase-free water (Qiagen). Reaction conditions were: 25 °C for 10 min, 42 °C for 60 min and 70 °C for 10 min. cDNA was used immediately or stored at −80 °C.

RT-qPCR was carried out using the Lightcycler 480 II (Roche, West Sussex, UK). Primers for RT-qPCR were manually designed to be under 200 bp in length for the region of interest (Supplementary Table 1). Each reaction consisted of 1x SYBR Green I Master (Roche), 0.5 μM of each primer and was made up to a 10 μl reaction using 1 μl cDNA and nuclease-free water (Qiagen). Reactions were: 95 °C for 10 min, 50 cycles of 95 °C for 10 s, 60 °C for 10 s and 72 °C for 10 s. Expression was normalised to *Actb* and relative expression was analysed via the $\Delta\Delta C_T$ method.

4.7. Glu-qPCR—quantification of 5hmC and 5mC

For germ cells, absolute levels of 5hmC at specific CpG sites (MspI restriction sites—CCGG) were quantified using the EpiJET 5-hmC Analysis Kit (Thermo Scientific) according to the manufacturer's instructions using approximately 400 oocytes or 250 ng of intact DNA (sperm), followed by RT-qPCR. For tissues, absolute levels of each DNA modification (C, mC and hmC) at specific DNA loci (HpaII/Msp1 restriction sites—CCGG) were quantified using the EpiMark 5hmC and 5mC Analysis Kit (New England Biolabs) according to the manufacturer's instructions. Absolute levels of each modification, spanning a single CCGG site, were then calculated based on the Cq (threshold) values, normalised to a mock digestion (containing no enzyme).

4.8. Bisulfite methylation analysis

Pyrosequencing was carried out as previously described [16] using pre-designed assays (Qiagen), excepting the intragenic assays for *Adcy6*, *Crocc* and *Efna2* which were designed in-house using PyroMark Assay Design Software 2.0 (Supplementary Table 1). Bisulfite-treated DNA was PCR amplified in 1x Buffer, 0.4 μM dNTPs, 1 μM primers and $MgCl_2$ at an optimal concentration to the primer set and 0.01U Taq DNA Polymerase (all Invitrogen).

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ygeno.2014.08.013.

Acknowledgments

Work in the Walsh lab is supported by grants from the MRC (MR/J007773/1) and the HSC R&D office Endocrinology TRG. We are very grateful to Heather Broadley, Helen Wheadon, Fulvio Gandolfi and Hazel Mangan for tissues and help with analyses; to Masaki Okano for ES cells; to Heather McCook, Rhonda Black and Keith Thomas for technical support and to members of the Group for comments.

Author contributions

REI made initial observations of 5hmC, designed assays, carried out most experimental work, assembled figures and contributed to the manuscript; AT designed assays and carried out the initial transcriptional and bisulfite methylation analyses in 3T3 and ES cells; KON analysed human homologs, designed assays and contributed to project design; CPW designed and supervised the study, carried out bioinformatic analyses and wrote the manuscript.

References

[1] Y.F. He, B.Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C.X. Song, K. Zhang, C. He, G.L. Xu, Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA, Science 333 (2011) 1303–1307.
[2] H. Hashimoto, S. Hong, A.S. Bhagwat, X. Zhang, X. Cheng, Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation, Nucleic Acids Res. 40 (2012) 10203–10214.
[3] L. Shen, H. Wu, D. Diep, S. Yamaguchi, A.C. D'Alessio, H.L. Fung, K. Zhang, Y. Zhang, Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics, Cell 153 (2013) 692–706.
[4] G. Ficz, M.R. Branco, S. Seisenberger, F. Santos, F. Krueger, T.A. Hore, C.J. Marques, S. Andrews, W. Reik, Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation, Nature 473 (2011) 398–402.
[5] T.P. Gu, F. Guo, H. Yang, H.P. Wu, G.F. Xu, W. Liu, Z.G. Xie, L. Shi, X. He, S.G. Jin, K. Iqbal, Y.G. Shi, Z. Deng, P.E. Szabo, G.P. Pfeifer, J. Li, G.L. Xu, The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes, Nature 477 (2011) 606–610.
[6] R.R. Zhang, Q.Y. Cui, K. Murai, Y.C. Lim, Z.D. Smith, S. Jin, P. Ye, L. Rosa, Y.K. Lee, H.P. Wu, W. Liu, Z.M. Xu, L. Yang, Y.Q. Ding, F. Tang, A. Meissner, C. Ding, Y. Shi, G.L. Xu, Tet1 regulates adult hippocampal neurogenesis and cognition, Cell Stem Cell 13 (2013) 237–245.
[7] M.A. Hahn, R. Qiu, X. Wu, A.X. Li, H. Zhang, J. Wang, J. Jui, S.G. Jin, Y. Jiang, G.P. Pfeifer, Q. Lu, Dynamics of 5-hydroxymethylcytosine and chromatin marks in mammalian neurogenesis, Cell. Rep. 3 (2013) 291–300.
[8] R. Lister, E.A. Mukamel, J.R. Nery, M. Urich, C.A. Puddifoot, N.D. Johnson, J. Lucero, Y. Huang, A.J. Dwork, M.D. Schultz, M. Yu, J. Tonti-Filippini, H. Heyn, S. Hu, J.C. Wu, A. Rao, M. Esteller, C. He, F.G. Haghighi, T.J. Sejnowski, M.M. Behrens, J.R. Ecker, Global epigenomic reconfiguration during mammalian brain development, Science 341 (2013) 1237905.
[9] F. Neri, A. Krepelova, D. Incarnato, M. Maldotti, C. Parlato, F. Galvagni, F. Matarese, H. G. Stunnenberg, S. Oliviero, Dnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA methylation at gene bodies in ESCs, Cell 155 (2013) 121–134.
[10] H. Wu, V. Coskun, J. Tao, W. Xie, W. Ge, K. Yoshikawa, E. Li, Y. Zhang, Y.E. Sun, Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes, Science 329 (2010) 444–448.
[11] J.P. Thomson, P.J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A.R. Kerr, A. Deaton, R. Andrews, K.D. James, D.J. Turner, R. Illingworth, A. Bird, CpG islands influence chromatin structure via the CpG-binding protein Cfp1, Nature 464 (2010) 1082–1086.
[12] A.K. Maunakea, R.P. Nagarajan, M. Bilenky, T.J. Ballinger, C. D'Souza, S.D. Fouse, B. E. Johnson, C. Hong, C. Nielsen, Y. Zhao, G. Turecki, A. Delaney, R. Varhol, N. Thiessen, K. Shchors, V.M. Heine, D.H. Rowitch, X. Xing, C. Fiore, M. Schillebeeckx, S.J. Jones, D. Haussler, M.A. Marra, M. Hirst, T. Wang, J.F. Costello, Conserved role of intragenic DNA methylation in regulating alternative promoters, Nature 466 (2010) 253–257.
[13] R.S. Illingworth, U. Gruenewald-Schneider, S. Webb, A.R. Kerr, K.D. James, D.J. Turner, C. Smith, D.J. Harrison, R. Andrews, A.P. Bird, Orphan CpG islands identify numerous conserved promoters in the mammalian genome, PLoS Genet. 6 (2010) e1001134.
[14] W. Huang da, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (2009) 44–57.
[15] W. Huang da, B.T. Sherman, R.A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, Nucleic Acids Res. 37 (2009) 1–13.
[16] C.E. Rutledge, A. Thakur, K.M. O'Neill, R.E. Irwin, S. Sato, K. Hata, C.P. Walsh, Ontogeny, conservation and functional significance of maternally inherited DNA methylation at two classes of non-imprinted genes, Development 141 (2014) 1313–1323.
[17] S.A. Smallwood, S. Tomizawa, F. Krueger, N. Ruf, N. Carli, A. Segonds-Pichon, S. Sato, K. Hata, S.R. Andrews, G. Kelsey, Dynamic CpG island methylation landscape in oocytes and preimplantation embryos, Nat. Genet. 43 (2011) 811–814.
[18] H. Kobayashi, T. Sakurai, M. Imai, N. Takahashi, A. Fukuda, O. Yayoi, S. Sato, K. Nakabayashi, K. Hata, Y. Sotomaru, Y. Suzuki, T. Kono, Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks, PLoS Genet. 8 (2012) e1002440.
[19] A. Szwagierczak, S. Bultmann, C.S. Schmidt, F. Spada, H. Leonhardt, Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA, Nucleic Acids Res. 38 (2010) e181.
[20] E. Dobbin, P.M. Corrigan, C.P. Walsh, M.J. Welham, R.W. Freeburn, H. Wheadon, Tel/PDGFRbeta inhibits self-renewal and directs myelomonocytic differentiation of ES cells, Leuk. Res. 32 (2008) 1554–1564.
[21] S. Hitoshi, R.M. Seaberg, C. Koscik, T. Alexson, S. Kusunoki, I. Kanazawa, S. Tsuji, D. van der Kooy, Primitive neural stem cells from the mammalian epiblast differentiate

to definitive neural stem cells under the control of Notch signaling, Genes Dev. 18 (2004) 1806–1811.
[22] K. Woodfine, J.E. Huddleston, A. Murrell, Quantitative analysis of DNA methylation at all human imprinted regions reveals preservation of epigenetic stability in adult somatic tissue, Epigenetics Chromatin. 4 (2011) 1.
[23] S.G. Jin, X. Wu, A.X. Li, G.P. Pfeifer, Genomic mapping of 5-hydroxymethylcytosine in the human brain, Nucleic Acids Res. 39 (2011) 5015–5024.
[24] J.E. Loughery, P.D. Dunne, K.M. O'Neill, R.R. Meehan, J.R. McDaid, C.P. Walsh, DNMT1 deficiency triggers mismatch repair defects in human cells through depletion of repair protein levels in a process involving the DNA damage response, Hum. Mol. Genet. 20 (2011) 3241–3255.
[25] C.E. Nestor, R. Ottaviano, J. Reddington, D. Sproul, D. Reinhardt, D. Dunican, E. Katz, J. M. Dixon, D.J. Harrison, R.R. Meehan, Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes, Genome Res. 22 (2012) 467–477.
[26] S.A. Smallwood, G. Kelsey, De novo DNA methylation: a germ cell perspective, Trends Genet. 28 (2012) 33–42.
[27] B. Giardine, C. Riemer, R.C. Hardison, R. Burhans, L. Elnitski, P. Shah, Y. Zhang, D. Blankenberg, I. Albert, J. Taylor, W. Miller, W.J. Kent, A. Nekrutenko, Galaxy: a platform for interactive large-scale genome analysis, Genome Res. 15 (2005) 1451–1455.
[28] J.Y. Li, D.J. Lees-Murdock, G.L. Xu, C.P. Walsh, Timing of establishment of paternal methylation imprints in the mouse, Genomics 84 (2004) 952–960.
[29] D.J. Lees-Murdock, H.T. Lau, D.H. Castrillon, M. De Felici, C.P. Walsh, DNA methyltransferase loading, but not de novo methylation, is an oocyte-autonomous process stimulated by SCF signalling, Dev. Biol. 321 (2008) 238–250.